Molecular Ecology (2006) **15**, 731–738 doi: 10.1111/j.1365-294X.2006.02813.x

Extensive intraspecific polymorphism detected by SSCP at the nuclear C-*mos* gene in the endemic Iberian lizard *Lacerta schreiberi*

RAQUEL GODINHO,* VERA DOMINGUES,*‡ EDUARDO G. CRESPO† and NUNO FERRAND*§
*CIBIO — Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661 Vairão, Portugal, †Centro de Biologia Ambiental e Departamento de Biologia Animal, Faculdade de Ciências da Universidade de Lisboa. Campo Grande, 1749-016 Lisboa, Portugal, ‡Unidade de Investigacão em Eco-etologia. I.S.P.A., R. Jardim do Tabaco 34, 1149-041 Lisboa, Portugal, §Departamento de Zoologia e Antropologia, Faculdade de Ciências, Universidade do Porto, Praça Gomes Teixeira. 4099-002 Porto, Portugal

Abstract

C-*mos* is a highly conserved intronless gene that has proved useful in the analysis of ancient phylogenetic relationships within vertebrates. We selected the Iberian endemic Schreiber's green lizard (*Lacerta schreiberi*) that persisted in allopatric refugia since the late Pliocene to investigate the utility of the C-*mos* nuclear gene for intraspecific phylogeographic studies. Our combination of DNA sequencing with the high resolving power of single-strand conformational polymorphism (SSCP) effectively discriminated four common alleles showing strong population structuring (F_{ST} = 0.46). In addition, reconstruction of allele phylogenetic relationships further improved our understanding of C-*mos* spatial patterns of variation and allowed a comparison with previously described mitochondrial DNA data. Finally, limited sequencing of an extended C-*mos* fragment in six additional *Lacerta* species showed extensive polymorphism, to our knowledge representing a rare example of variation in a highly conserved nuclear gene.

Keywords: intraspecific polymorphism, *Lacerta schreiberi*, nuclear gene, C-*mos*, SSCP

Received 24 August 2005; revision accepted 20 October 2005

Introduction

Reconstructing patterns of genetic differentiation and past colonization routes of organisms has been simplified in the last years with the progressive improvement of molecular techniques (Avise 2000). Currently, phylogeographic studies are generalized and most of them involve analyses of animal mitochondrial DNA (mtDNA). The peculiarities of mtDNA inheritance mode (maternal and nonrecombinant) and the rapid evolution of its sequences are at the origin of this success since they often provide multiple haplotypes that can be ordered phylogenetically within a species (Avise 2000). Nevertheless, the matrilineal phylogeny registered by this molecule represents only a very small fraction of the total historical record within a sexual organismal pedigree (Avise 1998; Zhang & Hewitt 2003), as well as a sex-biased look at this record — dispersal and gene flow are highly asymmetric by gender in many species, with females commonly philopatric to natal sites. While the use of autosomal loci for phylogeographic purposes is still dominated by relevant challenges that include (i) the identification of genomic regions with suitable rapid evolution, (ii) the possibility of intragenic recombination, and (iii) the technical problems of isolating haplotypes one at a time from diploid DNA (Zhang & Hewitt 2003), the important limitations of mtDNA-based phylogeography have emphasized the need of acquiring information from multiple nuclear loci before drawing firm phylogeographic conclusions for any species.

Mos (E.C. 2.7.1.-) is a serine/threonine kinase with a mass of approximately 39 kDa that result from the expression of the nuclear C-*mos* proto-oncogene. This kinase is known to be a key regulator of the meiotic process of vertebrate oocytes (Sagata *et al*. 1988; Gebauer & Richter 1997). Therefore, the C-*mos* gene is expressed at high levels in germ cells and at very low levels, if any, in somatic cells. When expressed in somatic cells, however, the protein

Correspondence: Raquel Godinho, Fax: +351 252 661780; E-mail: rgodinho@mail.icav.up.pt

product can readily induce oncogenic transformation of the cells (see review by Sagata 1997). In the genomes of species ranging from amphibians to humans, C-*mos* is present as a single coding exon of approximately 2 kb (Yew *et al*. 1993), making it an attractive tool in phylogenetic studies. To date, C-*mos* has been almost exclusively used as a molecular marker in deep phylogenetic studies relating orders (Cooper & Penny 1997; Hedges & Poling 1999; García-Moreno *et al*. 2003) or families (Saint *et al*. 1998; Harris *et al*. 1999; Hughes & Baker 1999; Barker *et al*. 2001). More recently, however, the general belief that C-*mos* is a slow-evolving gene, coherent with its pivotal role in cell cycle, was re-evaluated when substantial progress in resolving relationships of passerine genera was provided (Lovette & Bermingham 2000).

It is well known that certain regions of the Iberian Peninsula were Pleistocene refugia for many plants and animals, even during the most extreme glacial periods. For species that were originally from Iberia, the glacial times, together with the geographical and topographical heterogeneity of the peninsula, favoured repeated fragmentation of initial distribution areas over time, thus providing opportunities for differentiation (Blondel & Aronson 1999). This fact promoted regional differentiation at various levels of organization, ranging from the more simple genetic structuring of populations to the more complex processes of subspeciation, or even speciation. Examples of this have been described for a variety of organisms that include the Mediterranean ragwort, *Senecio gallicus* (Comes & Abbott 1998), the European rabbit, *Oryctolagus cuniculus* (Branco *et al*. 2000, 2002), the golden-striped salamander, *Chioglossa lusitanica* (Alexandrino *et al*. 2000, 2002), and the Schreiber's green lizard, *Lacerta schreiberi* (Godinho *et al*. 2001; Paulo *et al*. 2001, 2002). This last species is endemic to Iberia with a distribution restricted to the northwestern part of this peninsula and with several isolated populations in southern mountains where is generally found near water courses in Atlantic forest areas with high precipitation and mild winters. Preliminary work with *L. schreiberi* using mtDNA cytochrome *b* revealed high levels of phylogeographic structuring with two very different allopatric lineages exhibiting sequence divergence ranging from 4.1% to 7.2% (Paulo *et al*. 2001). This deep mtDNA phylogenetic disjunction distinguishes Portuguese and northern Spain populations from those in the Spanish central system and probably correspond to a long isolation period that started in the late Pliocene (Paulo *et al*. 2001). Despite relevant differences in their effective population size and the stochastic variance associated to any sampling of independent gene trees (Hudson & Turelli 2003), it is thus to be expected that to this depth of mtDNA divergence could correspond a genetic imprint in the nuclear genome of this species, including those genes that were described as very conservative.

In this study, we describe the utility of the C-*mos* gene for a phylogeographic study using *L. schreiberi* as a model. We document the spatial distribution of C-*mos* variability in 19 populations sampled throughout the entire species' range using single-strand conformational polymorphism (SSCP) and compare the phylogeographic pattern generated with these data with that from both mtDNA and protein polymorphisms. Additionally, we sequenced some individuals of other *Lacerta* species (two to three per species from different localities) to better evaluate the relevance of our findings and to reconstruct a phylogeny of the alleles found for *L. schreiberi*.

Materials and methods

Sampling and DNA extraction

Lacerta schreiberi samples were collected in 19 localities across the species' distribution area (Fig. 1), representing a total of 414 samples. Animals were released at the capture site. DNA extraction was performed following the Sambrook *et al*. (1989) protocol.

PCR amplification of C-mos *and sequencing*

Polymerase chain reaction (PCR) primers were designed by eye using sequences from three reptile families (Scincidae, Iguanidae and Agamidae) and the chicken, all available in GenBank. The gene regions for the primer design were selected for their conserved amino acid sequences and relatively low required degeneracy. Primer lengths were optimized to achieve approximately equal

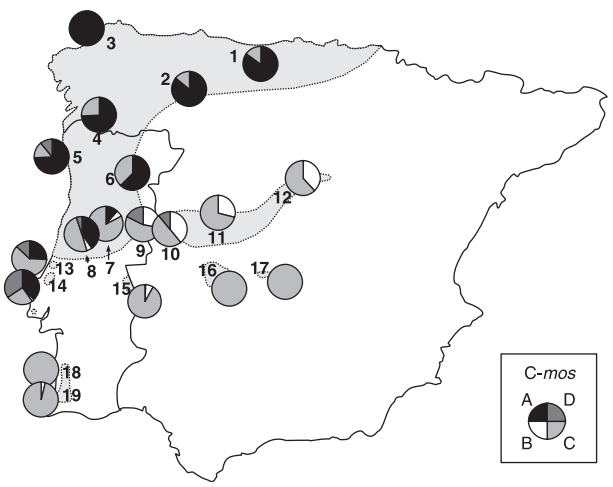

Fig. 1 Sampling locations of *Lacerta schreiberi* in the Iberian Peninsula. 1, Astúrias; 2, Ancares; 3, El Ferrol; 4, Gerês; 5, Gião; 6, Montemuro; 7, Estrela; 8, Lousã; 9, Malcata; 10, Gata; 11, Béjar; 12, Guadarrama; 13, C. Rainha; 14, Montejunto; 15, S. Mamede; 16, Guadalupe; 17, Toledo; 18, Cercal; 19, Monchique. The shaded area of the map represents the current distribution of the species. Pie charts represent the approximate frequencies of C-*mos* alleles.

estimates for annealing temperature. The forward primer, Mos-F, was a 29-mer with sequence 5′-CTCTGGKGGC-TTTGGKKCTGTSTACAAGG-3′ [443], and the reverse primer, Mos-R, was a 27-mer with sequence 5′-GGTGAT-GGCAAANGAGTAGATGTCTGC-3′ [1044]. Numbers in brackets after the primer refer to the 5′ position of the primer, as localized on the nucleotide sequence of the human C-*mos* gene (J00119). PCR conditions were optimized using a Robocycler Gradient 96 (Stratagene) and the thermal profile was defined as 3 min at 95 °C followed by 30 cycles at 95 °C for 1 min, 60 °C for 1 min, and 72 °C for 1 min. A 5-min extension at 72 °C was added on the end of the reaction. PCR amplifications were performed in 25-μL total volume, with 0.08 U of *Taq* polymerase (Promega), 12 pmol of each primer, 3 mM $MgCl_2$ and 0.2 mM of each dNTP. The resulting single PCR bands of 578 bp were sequenced for both strands with the PCR primers on an ABI PRISM 310 Genetic Analyser (PE Applied Biosystems) following the ABI PRISM BigDye Terminator Cycle sequencing protocols. Centrisep spin columns (Princeton Separations Inc.) were used for excess dye extraction. Sequencing was performed initially in a total of 10 samples of *L. schreiberi* from different populations. We also sequenced two to three individuals of other seven *Lacerta* species from different localities to look for polymorphic positions.

Detection of haplotypes by PCR–SSCP

Using specific sequences of *L. schreiberi*, we designed PCR–SSCP primers to target a fragment of 170 bp in the 3′ region. The selected forward and reverse primers were MosLs-F 5′-GTGCATCTGGATTTGAAACCTGCCAAC-3′ and MosLs-R 5′-CTCTCACCTTTAAGGAGTTCAGGA-GCAC-3′, respectively. PCR conditions were 3 min at 95 °C and 35 cycles of 1 min at 95 °C, 30 s at 64 °C and 30 s at 72 °C, followed by a 2-min extension at 72 °C. The reaction mixture was made for 10 μL total volume with 0.1 U of *Taq* polymerase (Promega), 5 pmol of each primer, 3 mM $MgCl_2$ and 0.2 mM of each dNTP. For PCR checking prior to SSCP analysis, amplifications were electrophoresed in 2% agarose gels. Preliminary SSCP tests were performed with samples known to carry different sequences. Discrimination of conformers using the SSCP technique was accomplished as follows: 1 μL of amplified DNA was mixed with 5 μL of denaturing loading buffer (95% deionized formamide, 10 mM NaOH, 0.01% bromophenol blue and 0.01% xylene cyanol). Samples were denatured for 5 min at 95 °C and kept on ice until gel loading. Five microlitre of sample was run in a 12% polyacrylamide gel (49:1 acrylamide : methylbisacrylamide) with 1 × TBE buffer on a vertical electrophoresis system (BIORAD Protean II). The electrophoresis was performed at a constant voltage of 250 V and constant temperature (12 °C) for 15 h. Routine SSCP separations always included previously typed samples that served as standards to ensure correct genotype scoring. The results were visualized by silver staining. Sequencing of four homozygous individuals for each allele collected in different populations was performed as described above in order to confirm identity of conformers.

Data analysis

Allele frequencies and measures of genetic variation were calculated using the GENETIX software, version 4.01 (Belkhir *et al*. 2000). Exact tests were used to analyse Hardy–Weinberg equilibrium with the help of the package GENEPOP (Raymond & Rousset 1995). Estimates of nucleotide diversity (π) were obtained with the DNASP 3.51 software (Rozas & Rozas 1999) while C-*mos* nucleotide contents and the transition/tranversion ratio were calculated using the MEGA 2.1 software (Kumar *et al*. 2000). The different C-*mos* haplotypes observed in *L. schreiberi* and related outgroup species were used to construct a neighbour-joining tree also with the help of MEGA 2.1 software. To compare levels of C-*mos* and mitochondrial differentiation we followed the methodology described in Lovette & Bermingham (2000). Nested clade phylogeographical analysis (NCPA) was applied to the C-*mos* data set as recently summarized by Templeton (2004).

Results and discussion

Partial C-*mos* gene sequences of 522 bp were obtained for eight lacertid species (six species included in the *Lacerta sensu stricto* group, *L. schreiberi*, *L. bilineata*, *L. viridis*, *L. agilis*, *L. trilineata* and *L. pamphylica*, and two outgroup taxa, *L. lepida* and *L. monticola*) representing 22 different sequences (Table 1). We found a total of 40 segregating sites from which 20 were parsimony informative and five out of the eight species exhibited nonsynonymous polymorphisms. Variable positions are distributed uniformly along the region studied and the transition/transversion substitution ratio was 2.3, which is in agreement with values reported for other reptile and bird species (Saint *et al*. 1998; Lovette & Bermingham 2000). Observed base composition averaged across the *Lacerta* species included in this study reflected a slight enrichment in adenines (T = 24.8%, C = 21.8%, A = 30.6%, G = 22.8%). No indels were found in any of these sequences when aligned among them. In addition, no indels were also found when sequences of *L. schreiberi* were compared with other species within the family Lacertidae. Nevertheless, 21 bp [1271] plus 3 bp [1461] corresponding to seven and one codons, respectively, were missing in all sequences when compared to the chicken sequence (numbers in brackets correspond to the first missing base in the *Gallus* C-*mos* nucleotide sequence).

Table 1 Variable positions found in a 522-bp fragment of the C-*mos* gene in *Lacerta schreiberi* and an additional set of five related species (*L. bilineata*, *L. viridis*, *L. agilis*, *L. trilineata* and *L. pamphylica*) and two outgroups (*L. lepida* and *L. monticola*). The fragment (170 bp) corresponding to the SSCP analysis is indicated on top of the table and relevant polymorphic positions in *L. schreiberi* are shaded. Boxes represent nucleotide variation that originates amino acid substitutions. Position 1 corresponds to position 472 of human C-*mos* gene (J00119)

SSCP fragment (positions 318–509)

		3	14	33	53	56	58	66	68	69	75	90	106	123	132	141	153	163	180	183	201	203	207	213	251	267	283	288	289	294	318	369	384	393	399	417	427	436	437	443	465	504	509
L. schreiberi	Allele C	A	C	G	A	G	T	A	A	G	C	A	G	C	A	C	T	G	A	A	T	G	T	C	G	G	C	C	T	T	A	A	C	C	T	C	C	C	C	A	T	T	G
L. schreiberi	Allele C2	.	.	.	G	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
L. schreiberi	Allele C3	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
L. schreiberi	Allele A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.
L. schreiberi	Allele B	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	.	.
L. schreiberi	Allele B2	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	.	.
L. schreiberi	Allele D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.
L. bilineata	Italy	.	.	A	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	G	.	.	.	.	.	.	G	.	.	C	.	.
L. bilineata	Spain	.	.	A	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	G	C	.	.	.	.	.	G	.	.	C	.	.
L. bilineata	Sicily	.	.	A	.	.	.	.	G	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	T	.	T	.	.	.	G	.	.	.	.	.	.	G	.	.	C	.	.
L. viridis	Croatia	.	.	A	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	G	.	.	.	.	.	.	G	.	.	C	.	.
L. viridis	Hungary	.	.	A	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	G	.	T	.	.	A	.	G	.	.	C	.	.
L. viridis	Greece	.	.	A	.	.	C	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	G	.	.	T	.	.	.	G	.	.	C	.	.
L. agilis	Austria	G	.	.	.	.	.	.	G	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	.	.	.	.	.	.	.	.	.	C	.	.
L. agilis	Holland	G	.	.	.	.	.	.	G	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	C	A	.
L. trilineata	Bosnia	G	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	G	.	.	C	.	.
L. trilineata	Krete	G	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	G	T	.	C	.	A
L. pamphylica	Turkey	G	.	.	.	A	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	G	.	.	C	.	.
L. pamphylica	Turkey	G	.	.	.	A	.	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	G	.	.	C	.	.
L. lepida	Portugal	.	A	.	.	.	.	.	G	.	.	G	.	G	.	T	.	A	.	.	.	.	.	.	.	A	.	A	.	C	G	.	.	.	.	.	.	G	.	.	G	.	.
L. lepida	Portugal	.	A	.	.	.	.	.	G	.	.	G	.	G	.	T	.	.	.	.	.	.	.	.	.	A	.	A	.	C	G	.	.	.	.	.	.	G	.	.	G	.	.
L. monticola	Portugal	G	.	.	.	.	.	.	G	.	T	.	.	G	G	.	.	.	C	C	.	.	C	G	.	.	.	.	.	C	G	.	.	.	.	.	G	G	.	.	G	.	.

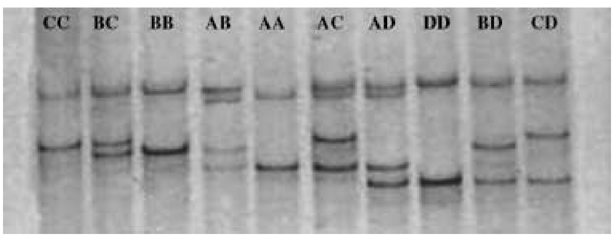

Fig. 2 Separation of allelic variants of the C-*mos locus* (170 bp) in *Lacerta schreiberi* by SSCP analysis on 12% polyacrylamide gels. Visualization was done by silver staining.

When the C-*mos* sequences obtained for the six species included in the *Lacerta sensu stricto* group were analysed with a sliding window of 100 bp and 10 bp steps (results not shown), two variable regions were found, one in the 5′ extreme and the other in the 3′ end of the amplified fragment. Due to a slightly higher number of polymorphic positions detected in the 3′ located variable region, we selected this portion of the C-*mos* gene for an extensive SSCP survey of 414 samples collected in 19 *L. schreiberi* populations encompassing the whole distribution area of this species. Our SSCP technique resolved a total of four alleles that were easily scored (A, B, C and D; Fig. 2), which were found in polymorphic frequencies ($p_i > 0.01$). The high resolving power of the SSCP technique was confirmed by sequencing four homozygous individuals for each SSCP allele which did not reveal additional variable positions. The phylogenetic relationships of the four SSCP alleles were easily determined by constructing a simple network and the ancestral haplotype was inferred from the sequences of related *Lacerta* species (Fig. 3). Two independent mutations in the ancestral haplotype are at the origin of alleles B (A→G transition) and C (C→T transition), while the two remnant alleles A and D derive from allele C by a single point mutation each (C→A transversion and T→C transition, respectively). The mutations originating alleles A and B implicate an amino acid substitution, in both cases modifying the net charge of the C-*mos* gene product (Fig. 3).

The C-*mos* allelic frequencies, together with the expected heterozygosity and nucleotide diversity (π), calculated for

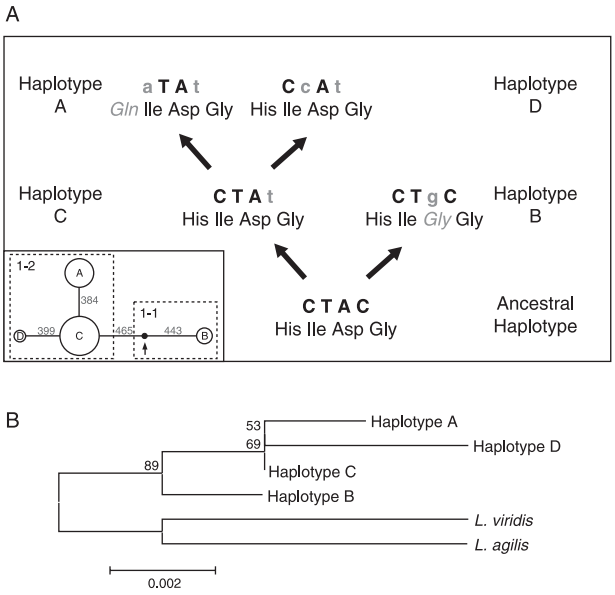

Fig. 3 (A) Phylogenetic relationships observed between the four SSCP C-*mos* haplotypes with the indication of nucleotide variable positions and corresponding amino acid substitutions. The ancestral allele was obtained by comparing C-*mos* sequences in *Lacerta schreiberi* with those from five closely related species. The inset shows the most parsimonious network where circle size is proportional to the frequency of each haplotype in the total sample. The root of the network is indicated by an arrow and corresponds to a missing haplotype (black circle). Numbers above lines correspond to the mutations that separate C-*mos* haplotypes. Nested clades designed for NCPA are represented with dashed lines. (B) Neighbour-joining tree relating *Lacerta schreiberi* C-*mos* haplotypes with two outgroup species. Bootstrap values (1000 replicates) are indicated.

the 19 *L. schreiberi* populations analysed in this study, are presented in Table 2. No deviations from Hardy–Weinberg expectations were detected over all populations. The most diverse populations were found along the Iberian central system, while a clear decrease in genetic diversity was detected in the northwestern region and in both the Portuguese southern isolates (Monchique and Cercal) and the Spanish central isolates (Guadalupe and Toledo). These results are in close agreement with a recent analysis of protein variation (Godinho *et al*. 2003) and also with the nucleotide diversity values reported for the mtDNA cytochrome *b* (Paulo *et al*. 2001).

Estimates of population differentiation revealed strong structuring (F_{ST} = 0.46) in the whole distribution area of *L. schreiberi*, but this value decreased dramatically when only five northwestern (F_{ST} = 0.06) or three Spanish central system (F_{ST} = 0.02) populations were analysed. A closer inspection of the allele frequency distribution in the Iberian Peninsula (Fig. 1 and Table 2) together with the analysis of phylogenetic relationships between alleles provide new insights into the evolution of the C-*mos* gene. Alleles B and C differ by two mutations and the missing haplotype corresponds precisely to the root of the network (Fig. 3). On the other hand, allele B is basically confined to the Iberian central system, exhibiting higher frequencies in the Spanish locations. It is thus probable that allele B originated in this region during the prolonged geographical isolation of corresponding populations and that only recently was able to disperse to the Portuguese mountains where it occurs in low frequency. Allele C is the most frequent and widespread allele and originated the derived alleles A and D by simple point mutations. Allele A characterizes all northwestern populations and probably marks the recent postglacial expansion of *L. schreiberi* into this region (Paulo *et al*. 2002), while allele D is restricted to the western Iberian mountains. This observation suggests that its origin may be related with the allopatric divergence of these populations, that was responsible for the emergence of a distinctive mtDNA coastal sublineage (Paulo *et al*. 2001) and a clear genetic imprint in several nuclear loci (Godinho *et al*. 2003 and unpublished results). NCPA of our C-*mos* data showed a clear lack of power due to the very simple haplotype tree analysed. Briefly, C-*mos* clade 1-1 (Fig. 3a) has no genetic variation and could not be included in the analysis, while clade 1-2 reflected long-distance colonization possibly coupled with subsequent fragmentation (results not shown).

Sequencing of a limited number of *L. schreiberi* samples for the 580-bp PCR-amplified fragment revealed three additional C-*mos* variants that can be considered subtypes of allele C (C2 and C3) and B (B2), all observed in the more diverse populations of central Iberian Peninsula (Table 1). Surprisingly, the sequencing of only two to three samples from related *Lacerta* species (including the more distant *L. lepida*, here used as an outgroup) showed a considerable number of variable positions in all of them, including multiple amino acid substitutions (Table 1). While a detailed investigation of the entire C-*mos* gene may eventually provide additional insights into the phylogenetic relationships and phylogeography of the lizard species included in the *Lacerta sensu stricto* group (Godinho *et al*. 2005), our present study combining limited DNA sequencing with the high resolving power of SSCP (Sunnucks *et al*. 2000) still captures the information provided by the C-*mos* gene and effectively demonstrates the utility of an additional nuclear marker to further improve our understanding of the evolutionary history of *L. schreiberi* in the Iberian Peninsula.

Since its first report as a good candidate gene for deep phylogenetic studies (Graybeal 1994), nucleotide sequences from the C-*mos* proto-oncogene have been mostly used in the reconstruction of relationships between distantly related vertebrate taxa (Cooper & Penny 1997; Saint *et al*. 1998; Harris *et al*. 1999; Hedges & Poling 1999; Hughes & Baker 1999; Barker *et al*. 2001; García-Moreno *et al*. 2003).

Table 2 Number of individuals sampled (*N*), allelic frequencies, number of alleles per population, expected heterozygosity (H_E) and nucleotide diversity (π) at the C-*mos* locus. The mtDNA lineages were from Paulo *et al*. (2001) with the following correspondence: A_1, coastal northern clade; A_2, coastal southern clade; B_1, inland northern clade; B_2, inland southern clade

			C-*mos* alleles								
	mtDNA lineage	*N*	A	B	C	D	No. of alleles	H_E	H_E proteins*	$\pi \pm$ SD	$\pi \pm$ SD mtDNA
Continuously distributed populations											
1. Asturias	A_1	13	0.846	—	0.154	—	2	0.260	0.06	0.00159 ± 0.00058	
2. Ancares	A_1	21	0.857	—	0.143	—	2	0.245	0.06	0.00148 ± 0.00046	
3. Ferrol	A_1	19	1.000	—	—	—	1	0.000	0.08	—	
4. Gerês	A_1	27	0.741	—	0.259	—	2	0.384	0.08	0.00230 ± 0.00034	
5. Gião	A_1	27	0.741	—	0.148	0.111	3	0.417	0.03	0.00467 ± 0.00103	0.00128 ± 0.00027
6. Montemuro	A_1	17	0.618	—	0.382	—	2	0.472	0.07	0.00286 ± 0.00026	
7. Estrela	A_1	26	0.115	0.058	0.827	—	3	0.300	0.09	0.00248 ± 0.00074	
8. Lousã	A_1	10	0.400	0.050	0.500	0.050	4	0.585	0.10	0.00520 ± 0.00125	
9. Malcata	A_1/B_1	25	—	0.280	0.540	0.180	3	0.598	0.09	0.00596 ± 0.00021	
10. Gata	B_1	22	—	0.386	0.500	0.114	3	0.588	0.08	0.00602 ± 0.00019	
11. Béjar	B_1	23	—	0.283	0.717	—	2	0.413	0.12	0.00488 ± 0.00070	0.00283 ± 0.00027
12. Guadarrama	B_1	21	—	0.381	0.619	—	2	0.472	0.08	0.00568 ± 0.00046	
Isolated populations											
13. S. Mamede	A_1	36	0.014	0.069	0.917	—	3	0.155	0.06	0.00171 ± 0.00062	
14. C. Rainha	A_2	27	0.259	—	0.611	0.130	3	0.543	0.09	0.00501 ± 0.00080	
15. Montejunto	A_2	24	0.375	0.021	0.250	0.354	4	0.671	0.08	0.00839 ± 0.00043	0.00295 ± 0.00051
16. Cercal	A_2	17	—	—	1.000	—	1	0.000	0.04	—	
17. Monchique	A_2	28	—	0.036	0.964	—	2	0.069	0.05	0.00083 ± 0.00054	
18. Guadalupe	B_2	13	—	—	1.000	—	1	0.000	0.03	—	0.00090 ± 0.00026
19. Toledo	B_2	18	—	—	1.000	—	1	0.000	0.00	—	

*Data from Godinho *et al*. (2003) and Godinho (2004); †data from Paulo *et al*. (2001).

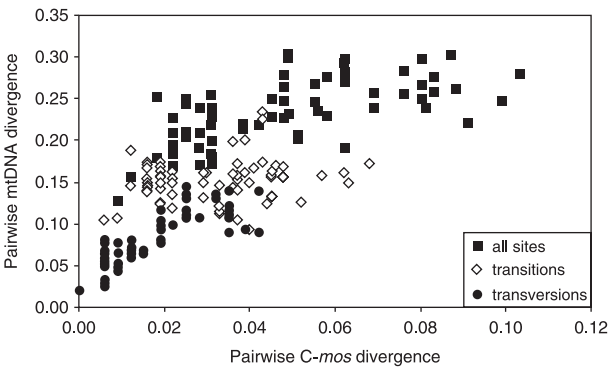

Fig. 4 Pairwise mitochondrial distances plotted against corresponding pairwise C-*mos* differentiation for 12 Lacertidae species.

Recently, however, Lovette & Bermingham (2000) convincingly demonstrated its utility in resolving phylogenetic relationships at intermediate levels of divergence in birds. In order to evaluate the generalization of their observation, we followed the methodology described in Lovette & Bermingham (2000) and plotted uncorrected mitochondrial (cytochrome *b*) distances against the corresponding gamma-corrected C-*mos* divergence values for 12 species within the family Lacertidae (data available from the authors upon request) (Fig. 4). Our results are remarkable in the sense that they fully confirmed the observations described for birds, suggesting that C-*mos* nucleotide substitutions accumulate at a rate similar to that of mitochondrial transversion substitutions also in lizards, thus providing useful, independent and complementary markers in phylogenetic studies of Lacertidae. At a lower taxonomic level, our detailed study of a small (170 bp) fragment together with a more limited analysis of an extended 580-bp fraction of the approximate 2 kb C-*mos* exon showed extensive intraspecific polymorphism, to our knowledge representing a rare example of variation in a highly conserved nuclear gene. Interestingly, two out of the four detected SSCP alleles implicated amino acid substitutions, resulting in two coding SNPs that can be used in future comparisons with noncoding SNPs and help in understanding the roles of population history and selection in shaping the genetic characteristics of this and other species.

Acknowledgements

We thank Paulo Alexandrino, James Harris and Jorge Rocha for their manuscript comments. We are also indebted to two anonymous referees for providing useful suggestions that improved a previous version of this manuscript and to James Harris for reviewing the English. We thank the Portuguese Institute of Nature Conservation for the collecting permits in Portugal and the Junta de Castilla y León, Junta de Extremadura, Junta de Comunidades de Castilla-La Mancha, Xunta de Galicia and Principado de Asturias for the collecting permits in Spain. We are grateful to all colleagues that helped in the field work, especially Cristina Luís. RG worked under a PhD grant (PRAXIS XXI/BD/16087/98) from Fundação para a Ciência e a Tecnologia (FCT), Portugal, and is presently supported by a post-doctoral grant (SFRH/BPD/12723/2003). NF was supported by a sabbatical grant (SFRH/BSAB/318/2002). This work was partially supported by FCT through POCTI/BIA/11174/98 and POCTI/BSE/40987/2001 grants.

References

Alexandrino J, Arntzen JW, Ferrand N (2002) Nested clade analysis and the genetic evidence for population expansion in the phylogeography of the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). *Heredity*, **88**, 66–74.

Alexandrino J, Froufe E, Arntzen JW, Ferrand N (2000) Genetic subdivision, glacial refugia and postglacial recolonization in the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). *Molecular Ecology*, **9**, 771–781.

Avise JC (1998) The history and purview of phylogeography: a personal reflection. *Molecular Ecology*, **7**, 371–379.

Avise JC (2000) *Phylogeography. The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts.

Barker FK, Barrowclough GF, Groth JG (2001) A phylogenetic hypothesis for passerine birds: taxonomic and biogeographic implications of an analysis of nuclear DNA sequence data. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **269**, 295–308.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2000) GENETIX *4.01, logiciel sous Windows™ pour la génétique des populations*. Laboratoire Génome et Populations, CNRS UPR 9060, Université de Montpellier II, Montpellier, France.

Blondel J, Aronson J (1999) *Biology and Wildlife of the Mediterranean Region*. Oxford University Press, New York.

Branco M, Ferrand N, Monnerot M (2000) Phylogeography of the European rabbit (*Oryctolagus cuniculus*) in the Iberian Peninsula inferred from RFLP analysis of the cytochrome *b* gene. *Heredity*, **85**, 307–317.

Branco M, Monnerot M, Ferrand N, Templeton AR (2002) Postglacial dispersal of the European rabbit (*Oryctolagus cuniculus*) on the Iberian Peninsula reconstructed from nested clade and mismatch analyses of mitochondrial DNA genetic variation. *Evolution*, **56**, 792–803.

Comes HP, Abbott RJ (1998) The relative importance of historical events and gene flow on the population structure of a Mediterranean ragwort, *Senecio gallicus* (Asteraceae). *Evolution*, **52**, 355–367.

Cooper A, Penny D (1997) Mass survival of birds across the Cretaceous–Tertiary boundary: molecular evidence. *Science*, **275**, 1109–1113.

García-Moreno J, Sorenson MD, Mindell DP (2003) Congruent avian phylogenies inferred from mitochondrial and nuclear DNA sequences. *Journal of Molecular Evolution*, **57**, 27–37.

Gebauer F, Richter JD (1997) Synthesis and function of Mos: the control switch of vertebrate oocyte meiosis. *BioEssays*, **19**, 23–28.

Godinho R (2004) *Evolutionary history of the Schreiber's green lizard*, Lacerta schreiberi: *Phylogenetic analysis, reconstruction of processes of population expansion, contraction and miscigenation, and integration in the biogeographic context of northwestern Iberian Peninsula*. PhD Thesis, University of Lisbon, Portugal.

Godinho R, Crespo EG, Ferrand N, Harris DJ (2005) Phylogeny and evolution of the green lizards *Lacerta* spp. (Squamata: Lacertidae) based on mitochondrial and nuclear DNA sequences. *Amphibia-Reptilia*, **26**, 271–285.

Godinho R, Ferrand N, Crespo EG (2001) Phylogeography of the Iberian Schreiber's green lizard (*Lacerta schreiberi*): preliminary data on mitochondrial and nuclear markers reveal discrepant patterns. *Mertensiella*, **13**, 33–39.

Godinho R, Paulo OS, Ferrand N *et al.* (2003) Major patterns of population differentiation in the Iberian Schreiber's green lizard (*Lacerta schreiberi*) inferred from protein polymorphism. *Herpetological Journal*, **13**, 35–42.

Graybeal A (1994) Evaluating the phylogenetic utility of genes: a search for genes informative about deep divergences among vertebrates. *Systematic Biology*, **43**, 174–193.

Harris DJ, Sinclair EO, Mercader NL, Marshall JC, Crandall KA (1999) Squamate relationships based on C-*mos* nuclear DNA sequences. *Herpetological Journal*, **9**, 147–151.

Hedges SB, Poling LL (1999) A molecular phylogeny of reptiles. *Science*, **283**, 998–1001.

Hudson RR, Turelli M (2003) Stochasticity overrules the 'three-times rule': genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA. *Evolution*, **57**, 182–190.

Hughes JM, Baker AJ (1999) Phylogenetic relationships of the enigmatic hoatzin (*Opisthocomus hoazin*) resolved using mitochondrial and nuclear gene sequences. *Molecular Biology and Evolution*, **16**, 1300–1307.

Kumar S, Tamura K, Jakobsen I, Nei M (2000) MEGA: *Molecular Evolutionary Genetics Analysis. version 2.1*. The Pennsylvania State University, University Park, Pennsylvania.

Lovette IJ, Bermingham E (2000) C-*mos* variation in songbirds: molecular evolution, phylogenetic implications, and comparisons with mitochondrial differentiation. *Molecular Biology and Evolution*, **17**, 1569–1577.

Paulo OS, Dias C, Bruford MW, Jordan WC, Nichols RA (2001) The persistence of Pliocene populations through the Pleistocene climatic cycles: evidence from the phylogeography of an Iberian lizard. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **268**, 1625–1630.

Paulo OS, Jordan WC, Bruford MW, Nichols RA (2002) Using nested clade analysis to assess the history of colonization and the persistence of populations of an Iberian lizard. *Molecular Ecology*, **11**, 809–819.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rozas J, Rozas R (1999) DNASP (version 3): an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*, **15**, 174–175.

Sagata N (1997) What does Mos do in oocytes and somatic cells? *BioEssays*, **19**, 13–20.

Sagata N, Oskarsson M, Copeland T, Brumbaugh J, Vande Woude GF (1988) Function of C-*mos* proto-oncogene product in meiotic maturation in *Xenopus* oocytes. *Nature*, **335**, 519–525.

Saint KM, Austin CC, Donnellan SC, Hutchinson MN (1998) C-*mos*, a nuclear marker useful for squamate phylogenetic analysis. *Molecular Phylogenetics and Evolution*, **10**, 259–263.

Sambrook E, Fritsch F, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sunnucks P, Wilson ACC, Beheregaray LB *et al.* (2000) SSCP is not so difficult: the application and utility of single-stranded conformation polymorphism in evolutionary biology and molecular ecology. *Molecular Ecology*, **9**, 1699–1710.

Templeton A (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology*, **13**, 789–809.

Yew N, Strobel M, Woude GFV (1993) Mos and the cell cycle: the molecular basis of the transformed phenotype. *Current Opinion in Genetics and Development*, **3**, 19–25.

Zhang D-X, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology*, **12**, 563–584.

This paper is part of the PhD thesis by Raquel Godinho on the evolutionary history of the endemic Iberian lizard *Lacerta schreiberi*. Vera Domingues worked on the project for her graduate thesis in Biology and has recently started up a project on the molecular biogeography of tropical and subtropical fishes of the Azores. Eduardo G. Crespo, professor at the Lisbon University, has been involved in several projects related with the study of amphibians and reptiles. Nuno Ferrand heads the CIBIO and is interested in a variety of questions in evolutionary and conservation genetics.
